

# *DataTurbine and HPWREN*

---

**Paul Hubbard**  
**hubbard@sdsc.edu**  
**Cyberinfrastructure Lab for Environmental Observing  
Systems**  
**Science R&D**  
**SDSC/UCSD**



OPEN SOURCE DATA  TURBINE INITIATIVE  
*Empowering the Scientific Community with Streaming Data Middleware*



UC San Diego

SAN DIEGO SUPERCOMPUTER CENTER, UCSD

# *The Big Picture*

- **Our group does middleware/cyberinfrastructure for science. We're collaborating with HPWREN to improve our software and contribute to new science.**
- **We have a weather station & camera deployed at the northern gorge of SMER, with a small server in the machine closet there.**
- **HPWREN has allowed us to stress and test our software on a wireless network and found many bugs.**



OPEN SOURCE DATA  TURBINE INITIATIVE  
*Empowering the Scientific Community with Streaming Data Middleware*

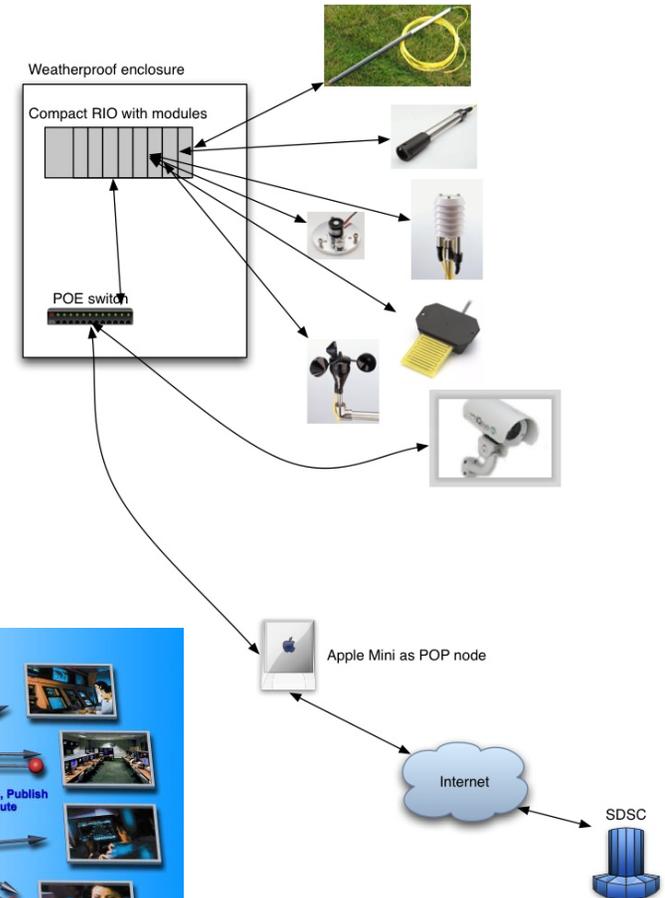


SAN DIEGO SUPERCOMPUTER CENTER, UCSD

# OPEN SOURCE DATA TURBINE INITIATIVE

Empowering the Scientific Community with Streaming Data Middleware

- DataTurbine is open-source software for managing real-time streaming data from sensors and instruments
- Supports heterogeneous instrument integration, data stream management, and distributed signal processing
- Active community of developers and users, numerous fielded applications in NSF and NASA projects
- [www.dataturbine.org](http://www.dataturbine.org)



OPEN SOURCE DATA TURBINE INITIATIVE

Empowering the Scientific Community with Streaming Data Middleware



SAN DIEGO SUPERCOMPUTER CENTER, UCSD

# The SMER deployment

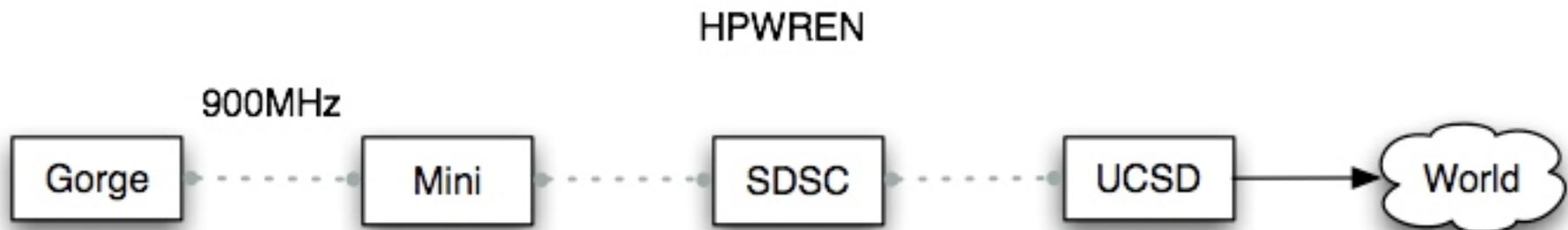


OPEN SOURCE DATA  TURBINE INITIATIVE  
*Empowering the Scientific Community with Streaming Data Middleware*



SAN DIEGO SUPERCOMPUTER CENTER, UCSD

# Dataflow and design



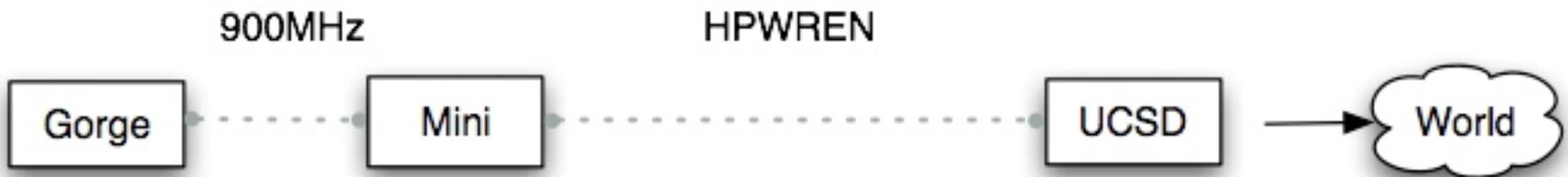
OPEN SOURCE DATA  TURBINE INITIATIVE  
*Empowering the Scientific Community with Streaming Data Middleware*



UC San Diego

SAN DIEGO SUPERCOMPUTER CENTER, UCSD

# Current system



# *Lessons learned and progress made*

- **Our Java code that reads data from the DAQ isn't robust enough against data loss and transmission errors.**
- **Our server mirroring code also failed several different ways and is still undergoing a lot of effort.**
- **Interop with current HPWREN software is a problem: cRIO lacks UDP multicast, and Java is not otherwise in use.**



# *More tips for other projects*

- **Network monitoring software is absolutely required**
  - ntop, SNMP-based such as Cacti, link monitors like smokeping, Probe-based (Inca), Monit, m/Monit.
- **Don't forget time sync (NTP)**
- **Use Power over Ethernet (PoE) if possible**
- **Low-level tools such as ngrep, socat, tcpdump and ntpq are often essential. Cultivate a network guru if possible.**
- **Web display of data is enormously valuable**



# *Current state and plans*

- **Currently looking for a science project as a driver (and of course funding)**
  - Fire detection?
  - Watershed monitoring?
  - Animal tracking?
  - Simple data sharing system for weather stations and video?
  - Web-based data distribution?
- **“We have the software, but we need the science.”**



## *...and of course*

- **Many thanks to HPWREN and its staff for the help in time, materials, bandwidth and patience!**



OPEN SOURCE DATA  TURBINE INITIATIVE  
*Empowering the Scientific Community with Streaming Data Middleware*



UC San Diego

SAN DIEGO SUPERCOMPUTER CENTER, UCSD